

## System and Method for Improving Transrating of MPEG-2 Video

### FIELD OF THE INVENTION

5

The present invention relates generally to systems and methods for transrating MPEG-2 video streams to meet the available bitrate of a transmission medium. More specifically the present invention relates to transrating the macroblocks of an MPEG-2 stream, which fall within a pan-scan area.

10

### BACKGROUND OF THE INVENTION

Throughout the disclosure and claims, we will be using the term MPEG (Motion Pictures Expert Group). MPEG is a generic reference to a family of international standards, which define how to encode visual and audio information in a digital compressed format.

15

MPEG is utilized in a wide variety of applications, including: DVD (Digital Video Discs) and DVB (Digital Video Broadcasting).

20

The MPEG standards specify exactly the format in which the compressed data is to be transmitted. A key feature of MPEG is that it can compress a video signal into a fraction of its original size. MPEG achieves a high compression for video by storing only the changes from one video frame to another, instead of each entire frame.

25

There are two major MPEG standards: MPEG-1 and MPEG-2. The most common implementations of the MPEG-1 standard provide video quality slightly below the quality of conventional VCR (Video Cassette Recorder) videos. MPEG-2

provides higher resolution, with full CD quality audio. This is sufficient for the major TV standards, including NTSC (National Standards Television Committee) and HDTV (High Definition Television).

5           Of the series of MPEG standards that describe and define the syntax for video broadcasting, the standard of relevance to the present invention is ISO/IEC IS 13818-2, ITU-T Recommendation the MPEG-2 standard, titled "Generic Coding of Moving Pictures and Associated Audio Information: Video.", which is incorporated herein by reference and is hereinafter referred to as "the MPEG-2 Standard"

10

Additional standards incorporated herein by reference are:

1)       MPEG-2 Systems Group. Information Technology- Generic Coding of Moving Pictures and Associated Audio: Part 1- Systems, ISO/IEC 13180-1  
15 International Standard 1995.

2)       MPEG-2 Video Group. Information Technology – Generic Coding of Moving Pictures and Associated Audio: Part 2 – Video ISO/IEC 13818-2  
International Standard, 1995.

20

If an MPEG stream is to be viewed immediately as it is received, the communication channel must have enough bit rate capacity to provide the series of pictures at a real-time rate. Bit rate is the number of digital bits which a communication channel can transmit per second. Alternatively, pictures can be  
25 encoded to a size suitable for a channel of a given bit rate. MPEG encoding allows the size of each picture to be adjusted by varying quality; smaller pictures may be achieved at the expense of lower quality. The objective of an MPEG encoding is to maximize quality for the available bit rate.

Consider a system in which high-quality MPEG video is to be played directly from a storage medium; DVD is an example. The video quality is high, as instantaneous bit rate is of relatively low concern. Consider now that the stored high-quality video is to be communicated across a channel of constrained bit rate such as a telephone line. Some device must re-encode each picture (with potential reduction of quality) so the sequence of pictures may be transmitted in real time within the available bit rate.

A transcoder is a device which converts an encoded bit stream of one bit rate to a lower bit rate, and changes the content of the encode bit stream. This device is useful when down-stream channels operate at a lower rate than the up-stream channel. An example would be the retrieval of data from a DVD drive at very high bit rate (and high quality) and transmission over phone line at a lower rate (and lower quality). The manner in which picture quality is traded-off for bits is a matter of transcoder design. Ideally, changes are made where they are least likely to be noticed by the viewer.

Traditional film mediums utilize the 35mm format. This format provides a 4:3 aspect ratio. An aspect ratio is the ratio of width to height of an image on a television, motion picture, or computer screen. As movie making technology developed, a wide screen format became more popular. This format is often referred to as "Cinemascope", "wide screen" or "letterbox". This format has an aspect ratio of 16:9.

When digital video is stored on a high-speed storage device, such as a DVD, the bit rate of the stream is often less important than the quality and inclusion of special features. One such special feature is the recording of a movie in a 16:9 aspect ratio, which doesn't use up all of the screen area of a standard TV (4:3 aspect ratio).

When viewed on a standard TV, a letterbox movie will have black bars at the top and bottom of the television screen, as a result of scaling down the image to fit the screen.

5 The MPEG-2 standard provides a means for embedding "pan-scan" information in a video bit stream. Pan-scan information defines which rectangular area of the stream is to be displayed for an aspect ratio different from the original recording.

10 U.S. Patent #5,638,130, discloses the use of pan-scan conversion, but after the MPEG-2 stream has been decoded. This requires that the MPEG-2 stream be decoded prior to selecting the pan-scan conversion. This does not address the issue of reducing bit rate requirements between the source and the end consumer.

15 There is thus a need for a transcoder capable of reducing the bit rate of an MPEG stream between the source of the MPEG stream and the final consumer. The present invention addresses this need.

### **SUMMARY OF THE INVENTION**

20 The present invention relates to a system and method for utilizing a transcoder to reduce the content of an MPEG-2 stream where applicable.

25 One aspect of the present invention is a method of determining if the content of a MPEG-2 source stream, may be reduced, the method having the steps of:

a) examining the source stream to determine if a sequence\_display\_extension follows the most recent sequence header and sequence extension;

b) confirming that horizontal\_size is greater than display\_horizontal\_size or that vertical\_size is greater than display\_vertical\_size; and

c) if steps a) and b) are met, reducing the content of the source stream to create a reformatted stream.

5

In another aspect of the present invention, there is provided a system for reducing the content of an MPEG-2 stream, the system containing:

a) a transcoder, the transcoder converting a source stream to a reformatted stream;

10 b) a transmitter connected to the transcoder, for transmitting the reformatted stream;

c) a receiver connected to the transmitter, for receiving the reformatted stream; and

15 d) a decoder connected to the receiver, for decoding the reformatted stream and providing as output a viewable stream.

In another aspect of the present invention, there is provided a computer readable medium containing instructions for reducing the content of an MPEG-2 source stream, said instructions performing the steps of:

20 a) examining the source stream to determine if a sequence\_display\_extension follows the most recent sequence header and sequence extension;

b) confirming that horizontal\_size is greater than display\_horizontal\_size or that vertical\_size is greater than display\_vertical\_size; and

25 c) if steps a) and b) are met, reducing the content of the source stream to create a reformatted stream.

In another aspect of the present invention there is provided a system for repositioning frames in an MPEG-2 stream said system including repositioning

means, the repositioning means utilizing pan-scan information to relocate a display rectangle to a reconstructed frame.

In another aspect of the present invention there is provided a video transcoder,  
5 the transcoder including a pan-scan module. The pan-scan module performing the steps of :

- 10 a) examining a source video stream to determine if a sequence\_display\_extension follows the most recent sequence header and sequence extension;
- b) confirming that horizontal\_size is greater than display\_horizontal\_size or that vertical\_size is greater than display\_vertical\_size; and
- c) if steps a) and b) are met, reducing the content of said source stream to create a reformatted stream.

15 In yet another aspect of the present invention there is provided a pan-scan module, the pan-scan module residing within a video transcoder, the module including:

- 20 a) a selector for examining a source stream to determine if the source stream may be reduced;
- b) a first calculator connected to the source stream to calculate the values of: width\_mb and height\_mb;
- c) a second calculator connected to the source stream for calculating the values  
25 of: top, bottom, left and right;
- d) a third calculator connected to the source stream for calculating the values of top\_mb, bottom\_mb, left\_mb and right\_mb; and
- e) a modifier taking as input the calculations performed by the first, second and third calculators to create a reformatted stream.

## **BRIEF DESCRIPTION OF THE DRAWINGS**

5           For a better understanding of the present invention, and to show more clearly how it may be carried into effect, reference will now be made, by way of example, to the accompanying drawings which aid in understanding an embodiment of the present invention and in which:

10           Figure 1 is a block diagram of a system utilizing the present invention;  
            Figure 2 is a block diagram of the syntax of an MPEG-2 video stream;  
            Figure 3 is a schematic diagram illustrating the repositioning of a frame;  
            Figure 4 is a flowchart of the source stream reformatting process;  
            Figures 5a and 5b are a block diagram of a transcoder; and  
15           Figure 6 is a block diagram illustrating the components of the Pan-scan module of Figure 5a.

## **DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS**

20           The transmission and reception of digital video requires complex hardware and software components. It is not the intent of this disclosure to address all such components but rather to address the specific areas within a digital video system in which the present invention may be utilized.

25           By way of introduction we refer first to Figure 1, a system utilizing the present invention, shown generally as 10. System 10 is an overview of a generic system in which the present invention may be utilized. System 10 comprises transcoder 12, transmitter 14, receiver 16 and decoder 18. Transcoder 12 accepts as input source stream 20. For the purpose of simplicity the reader may think of source stream 20 as

being a high bit rate MPEG-2 stream. However, it is not the intent of the inventors to restrict stream 20 to the format that is defined within the current MPEG-2 standard. Transcoder 12 receives source stream 20 and reformats it to a reformatted stream (not shown). The intent of creating a reformatted stream is to reduce the amount of data  
5 required to be sent to transmitter 14. The reformatted stream is sent to transmitter 14 via communications link 22. Transmitter 14 then forwards the reformatted stream to receiver 16 via communications link 24. Receiver 16 then forwards the reformatted stream to decoder 18 via communications link 26. Decoder 18 accepts the reformatted stream and decodes it to output a viewable stream 28 to the end user,  
10 which in this example would be one or more individuals viewing a video sequence on a display device, such as a television or computer monitor.

As one skilled in the art will recognize, communications links 22, 24, and 26 may take on a variety of forms such as, satellite transmission, wireless transmission,  
15 coaxial cable, twisted pair, internal computer bus, or any other form of transmitting data.

An MPEG stream is essentially a series of pictures taken at closely spaced time intervals. In the MPEG standards a picture is referred to as a "frame". This is  
20 terminology we will use from now on.

The basic building block of an MPEG frame is a macroblock. A macroblock consists of a 16x16 array of luminence (grayscale) pixels together with two 8x8 blocks for chrominance (colour) pixels. It is not the intent of this disclosure to discuss  
25 the technical details of the information in a macroblock as it is well known to those skilled in the art and well defined in the MPEG-2 standard. At the introductory level, one may consider a macroblock to be a digital equivalent of a portion of a traditional analog RGB (red, green, blue) television signal.



Referring now to Figure 2, a block diagram of the syntax of an MPEG-2 video stream is shown generally as 40. The syntax of steam 40 is described in detail in the MPEG-2 standard. The components of stream 40 of relevance to the present invention are: sequence header 42, sequence extension 44 and extension and user data 46.

Sequence header 42 contains the fields illustrated in Table 1.

**Table 1. Sequence Header (42)**

| Field Name            | Field Size      |
|-----------------------|-----------------|
| sequence_header_code  | 32 bits         |
| horizontal_size_value | 12 bits         |
| vertical_size_value   | 12 bits         |
| other                 | 36 to 1060 bits |

Sequence extension 44 contains the fields illustrated in Table 2.

**Table 2. Sequence Extension (44)**

| Field Name                      | Field Size |
|---------------------------------|------------|
| extension_start_code            | 32 bits    |
| extension_start_code_identifier | 4 bits     |
| profile_and_level_indication    | 8 bits     |
| progressive_sequence            | 1 bit      |
| chroma_format                   | 2 bits     |

|                           |        |
|---------------------------|--------|
| horizontal_size_extension | 2 bits |
| vertical_size_extension   | 2 bits |
| other                     | ...    |

As shown in Table 2, sequence extension 44 contains an “extension\_start\_code\_idenfifer” field. If this field contains the binary value of 0010 then sequence extension 44 defines a “sequence\_display\_extension” as illustrated in Table 3. The sequence\_display\_extension is contained within Extension and User data block 48 of Figure 2.

10 **Table 3. sequence\_display\_extension**

| Field Name                      | Field Size    |
|---------------------------------|---------------|
| extension_start_code_identifier | 4 bits (0010) |
| video_format                    | 3 bits        |
| colour_description              | 1 bit         |
| colour information              | 0 to 24 bits  |
| display_horizontal_size         | 14 bits       |
| marker_bit                      | 1 bit         |
| display_vertical_size           | 14 bits       |

As shown in Table 2, sequence extension 44 contains an “extension\_start\_code\_idenfifer” field. If this field contains the binary value of 0111 then the sequence extension defines a “picture\_display\_extension” as illustrated in Table 4. The picture\_display\_extension is contained within Extension and User data block 46 of Figure 2.

09905600.071301  
TDEP20.00950560

**Table 4. picture\_display\_extension**

| Field Name                      | Field Size    |
|---------------------------------|---------------|
| extension_start_code identifier | 4 bits (0111) |
| frame_centre_horizontal_offset  | 16 bits       |
| marker_bit                      | 1 bit         |
| frame_centre_vertical_offset    | 16 bits       |
| marker_bit                      | 1 bit         |

5       The block of contiguous fields in Table 4 comprising:  
frame\_centre\_horizontal\_offset, marker\_bit, frame\_centre\_vertical\_offset and  
marker\_bit, may occur from one to three times. The number of occurrences is based  
upon the value of “number\_of\_frame\_centre\_offsets” which is defined in the MPEG-  
2 standard.

10

The fundamental unit of measurement in repositioning a frame is the  
macroblock unit, which is a single row of a 16x16 luminance macroblock.

15       Within sequence\_header 42 (Figure 1, Table 1), are contained the 12-bit fields  
of horizontal\_size\_value and vertical\_size\_value. Within sequence\_extension 44  
(Figure2, Table 2), the 2-bit horizontal\_size\_extension and vertical\_size\_extension  
fields are found. The dimensions of a frame are defined by horizontal\_size and  
vertical\_size. The horizontal\_size is a 14-bit unsigned integer made up from  
horizontal\_size\_extension (making the most significant 2 bits) and  
20 horizontal\_size\_value (making the least significant 12 bits). Similarly, vertical\_size is  
a 14-bit unsigned integer made up from vertical\_size\_extension (making the most  
significant 2 bits) and vertical\_size\_value (making the least significant 12 bits).

10090500"071301  
T0ET20"00950660

In calculating the repositioning of a frame, we will make use of the functions floor() and ceil(). These functions are well known to those skilled in the art. The function floor() provides a rounding method that rounds toward zero by truncating. For example, 2.5 rounds to 2 and -2.5 rounds to -2. The function ceil() is a rounding method that rounds away from zero by truncating. For example, 2.5 rounds to 3 and -2.5 rounds to -3.

The width of each frame in macroblock units is:

10      width\_mb = ceil(horizontal\_size/16).

15      The height of each frame in macroblock units depends on whether the sequence of frames is interlaced or progressive. MPEG-2 supports two scanning methods, one is interlaced scanning and the other is progressive scanning. Interlaced scanning scans odd lines of a frame as one field, and even lines as another field. Progressive scanning scans the consecutive lines in sequential order.

In interlaced frames, frames may then be coded as either a frame picture or as two separately coded field pictures.

20      The 1-bit value progressive\_sequence is found within sequence\_extension 44 (Figure 2, Table 2). If progressive\_sequence contains a value of "1", the height of each picture in macroblock units is;

25      height\_mb = ceil(vertical\_size/16).

If progressive\_sequence contains a value of "0", either frame or field pictures may occur: the height of frame pictures in macroblock units is;

height\_mb = 2\*ceil(vertical\_size/32),

and the height of field pictures in macroblock units is

$\text{height\_mb} = \text{ceil}(\text{vertical\_size}/32).$

5

Each picture, be it a frame picture or a field picture may have associated with it a `picture_display_extension` (Table 4). Of interest to us are the fields: `frame_centre_horizontal_offset` and `frame_centre_vertical_offset`. The values in these fields are each 16-bit signed integer quantities representing the pan-scan offset in units of 1/16th of a sample. For `frame_centre_horizontal_offset`, a positive value indicates that the center of the reconstructed frame lies to the right of the center of the display rectangle. For `frame_centre_vertical_offset`, a positive value indicates that the center of the reconstructed frame lies below the center of the display.

10

15

By way of explanation of the above terminology, we now refer to Figure 3, a schematic diagram illustrating the repositioning of a frame. Display rectangle 50 contains the frame to be repositioned in reconstructed frame 52. In order to best position display rectangle 50 within reconstructed frame 52 the present invention makes use of `frame_centre_horizontal_offset` 54 and `frame_centre_vertical_offset` 56. The values of these fields are contained within the `picture_display_extension` as shown in Table 4. As shown in Figure 3, the centre of display rectangle 50 is illustrated as centre 58. Based upon the values of `frame_centre_horizontal_offset` 54 and `frame_centre_vertical_offset` 56, centre 58 becomes new centre 60, i.e. the centre of reconstructed frame 52.

20

25

The present invention will not modify source stream 20 (Figure 1) unless both of the following conditions are true:

1. A sequence\_display\_extension (Table 3) follows the most recent sequence header 42 and sequence extension 44. This is necessary so that the values of display\_horizontal\_size and display\_vertical\_size are known.

- 5    2. Either horizontal\_size is greater than display\_horizontal\_size or vertical\_size is greater than display\_vertical\_size. In either of these cases there may be macroblocks that lie outside of the viewable area.

10    If the above conditions are true, the present invention will modify source stream 20, executing as follows for each frame:

First, the display rectangle 50 is computed in the coordinates of the reconstructed frame 52 in units of 1/16 pixel:

15        top    =    (vertical\_size/2)\*16    -    frame\_centre\_vertical\_offset    -  
                 (display\_vertical\_size/2)\*16

          left   =   (horizontal\_size/2)\*16    -   frame\_centre\_horizontal\_offset    -  
                 (display\_horizontal\_size/2)\*16

20

          bottom = top + display\_vertical\_size\*16

          right = left + display\_horizontal\_size\*16

25    Next, the values of top, left, bottom and right are rounded off to whole macroblock units as shown by the following pseudo code.

**Value of top\_mb**

```

if top < zero then
    top_mb = 0
else
    if progressive_sequence == 1 then
5        top_mb = floor((top/16)/16)
    else
        if a field picture then
            top_mb = floor((top/16)/32)
        else
10        top_mb = 2*floor((top/16)/32)

```

#### Value of left\_mb

```

15 if left < zero then
    left_mb = 0
else
    left_mb = floor(left/256)

```

#### 20 Value of bottom\_mb

```

if progressive_sequence == 1
    bottom_mb = ceil((bottom/16)/16)
else
25 if a field picture
    bottom_mb = ceil((bottom/16)/32)
    else
        bottom_mb = 2*ceil((bottom/16)/32)

```

if bottom\_mb > (height\_mb-1)  
bottom\_mb = height\_mb -1

**Value of right\_mb**

5  
if ceil(right/256) < (width\_mb-1)  
right\_mb = ceil(right/256)  
else  
right\_mb = (width\_mb-1)

10

Thus the size of reconstructed frame 52 becomes:

new\_horizontal\_size = (right\_mb - left\_mb + 1) \* 16  
new\_vertical\_size = (bottom\_mb - top\_mb + 1) \* 16

15

To further aid the reader in understanding the above sequence of computations, we refer now to Figure 4, a flowchart of the source stream reformatting process, shown generally as 70. At step 72, source stream 20 is examined to determine if the reformatting conditions as described above have been met. If the conditions have been met the process moves to step 76 else the process terminates at step 74 as no reformatting is required. At step 76 the values of width\_mb and height\_mb are calculated as described above. These values define the width and height of display rectangle 50, before reconstruction, in macroblock units. At step 78 the values of top, bottom, left and right are calculated. These values define the position of the frame (display rectangle 50), within reconstructed frame 52. At step 80 the values of top, bottom, left and right are rounded off to whole macroblock units. At step 82, source stream 20 is modified to create a reformatted stream containing the new position of reconstructed frame 52. The reformatted stream is then output via communication link 22 (Figure 1).



The reformatted stream transmitted via communication links 22, 24 and 26 is a copy of source stream 20 (Figure 1), save for the following modifications.

- 5 1. In sequence\_header 42 (Figure 2, Table 1), horizontal\_size\_value takes the 12 least significant bits of new\_horizontal\_size and vertical\_size\_value takes the 12 least significant bits of new\_vertical\_size.
2. In sequence\_extension 44 (Figure 2, Table 2), horizontal\_size\_extension takes  
10 the 2 most significant bits of new\_horizontal\_size and vertical\_size\_extension takes the 2 most significant bits of new\_vertical\_size.
3. In sequence\_display\_extension (Table 3), display\_horizontal\_size takes the  
15 value of new\_horizontal\_size and display\_vertical\_size takes the value of new\_vertical\_size. If all of the other items in sequence\_display\_extension are default values then the sequence\_display\_extension does not need to be transmitted.
4. The picture\_display\_extension (Table 4) is not transmitted. It contained the  
20 frame\_center\_vertical\_offset and frame\_center\_horizontal\_offset values which are no longer required for reconstructed frame 52.
5. Macroblocks are not included if their horizontal position is less than left\_mb  
or greater than right\_mb or if their vertical position is less than top\_mb or greater than  
bottom\_mb. These are the macroblocks which lie outside of reconstructed frame 52.
- 25 6. Slices which now contain no macroblocks are not included. A slice is a consecutive series of macroblocks which are all located in the same horizontal row of a macroblocks.



originally applied to source stream 20. Motion Compensated Prediction module 100 accepts as input, motion compensation data contained in source stream 20 via link 93. Module 100 then provides picture difference information that is added at block 102 to produce decoded video stream 104. Module 100 also outputs via link 106  
5 characteristic information, such as motion estimation, as found in stream 20. This characteristic information is passed to the MCP module 126 of the encoder as is shown in Figure 5b.

With regard to the above description of Figure 5a, it is not the intent of the  
10 inventors to describe in detail the functionality of modules 94, 96, 98, 100 and 102. The functionality expected of such modules is defined in the MPEG-2 standard and numerous variations have been published and are well known to those skilled in the art.

Referring now to Figure 5b, decoded video stream 104 is input to module 112,  
15 which subtracts motion compensated prediction from the picture to form a "prediction error" picture. The prediction error picture is passed to Discrete Cosine Transform module 114 which transforms the picture and passes it to Quantiser module 116. Quantiser module 116 quantises the coefficients produced by module 114 and creates  
20 a new picture. The new picture may have to be recalculated based upon prediction error, thus the reason for the loop comprising blocks 120, 122 and 124 and 126. Motion Compensation Module 126 accepts as input characteristic information such as motion estimation, via link 106, as shown in both Figures 5a and 5b.

Once the picture has been successfully encoded it is passed to Variable Length  
25 Coder module 118. Module 118 then transmits the reformatted stream via communications link 22, to transmitter 14 (see Figure 1).

With regard to the above description of Figure 5b, it is not the intent of the inventors to describe in detail the functionality of the modules shown. The functionality expected of such modules is defined in the MPEG-2 standard and numerous variations are well known to those skilled in the art.

5

To further illustrate the functionality of the present invention we refer now to now to Figure 6, a block diagram illustrating the components of the Pan-scan module 92 of Figure 5a. The components of pan-scan module 92 provide the functionality illustrated in Figure 4 and as described herein. Selector 140 accepts as input source stream 20 and determines both criteria for modification have been met, namely:

10

1 A sequence\_display\_extension (Table 3) follows the most recent sequence header 42 and sequence extension 44. This is necessary so that the values of display\_horizontal\_size and display\_vertical\_size are known.

15

2 Either horizontal\_size is greater than display\_horizontal\_size or vertical\_size is greater than display\_vertical\_size. In either of these cases there may be macroblocks that lie outside of the viewable area.

20 Should these conditions be met, selector 140 passes the source stream 20 to first calculator 142. First calculator 14 calculates the value of width\_mb and height\_mb. First calculator 142 passes the source stream 20 to second calculator 144. Second calculator 144 calculates the values of: top, bottom, left and right. Second calculator passes the stream to third calculator 146 which calculates the values of: top\_mb, bottom\_mb, left\_mb and right\_mb. Third calculator passes the stream to modifier 25 148 which utilizes the values calculated by the first, second and third calculators (as described earlier with regard to Figure 1) to modify source stream 20 to create a reformatted stream which is output via communication link 22.

As one skilled in the art will recognize any of the components of pan-scan module 92 may be combined together and may indeed be a single hardware device or software procedure. The intent of the inventors when describing the functionality of pan-scan module 92 is to identify its functional steps, not to dictate a rigid structure..

5

The above disclosure describes how a transcoder can reduce the bit rate of a stream by transmitting only the visible area of the movie specified by pan-scan directives. Because the pan-scan conversion is performed prior to transmission, not immediately before display, the non-viewed portion of video does not need to be transmitted. Clearly the user loses the ability to select the letterbox format, but the reduction in bit rate and/or improvement in quality will justify this loss.

Although the invention has been described with reference to certain specific embodiments, various modifications thereof will be apparent to those skilled in the art without departing from the spirit and scope of the invention as outlined in the claims appended hereto.